PARAMETRIC MODELING AND DETECTION OF RIPPLE FIRED SIGNALS

Robert H. Shumway
Allan D.R. McQuarrie

Division of Statistics
University of California
Davis, CA 95616

AFTAC Contract F19628-95-K-0010

Abstract:

We develop a parametric statistical approach to detecting and estimating the delay and duration parameters of a ripple-fired signal. Such signals are generated by mining blasts with roughly equally spaced charges. Detection of the echo structure induced by ripple firing can serve as a basis for discriminating between mining blasts and nuclear explosions or earthquakes which, in principle, will not contain such echo effects.

The model assumes a noise corrupted signal that can be written as a convolution of source, path and instrument responses with a pulse sequence. The underlying responses are modeled by low order autoregressive moving average (ARMA) models that are consistent with conventional deterministic formulations of source theory and instrument response. The pulse sequence is modeled as a seasonal ARMA process, with the period corresponding to the firing delay and with a model order that is proportional to the signal duration. We use the cepstrum to suggest a range of values for the duration and delay parameters and then search all possible seasonal ARMA models within this range. The final model chosen is the minimizer of the corrected Akaike Information Theory Criterion (AIC_C).

Limited simulations are given to show that both the duration and delay can be estimated effectively with the seasonal ARMA search when the delays are equal and that reasonable estimators are available when the delays are variable. The methodology is applied also to the P phase of a single Scandinavian mining explosion where we obtain reasonable estimators for delay and duration under the assumption of ripple-firing.

Key Words: *Mining explosions, duration and delay, echo detection, cepstra, seasonal ARMA, model selection.*

OBJECTIVE

Regional seismic monitoring and discrimination capabilities that are desirable under a potential Comprehensive Test Ban Treaty (CTBT) can be improved by developing algorithms and new procedures for distinguishing between earthquakes, nuclear explosions and mining explosions of various kinds. Much effort in past discrimination studies has concentrated on extracting various features of the spectrum that are characteristic of earthquakes, nuclear explosions or mine blasts.

One particular spectral feature that characterizes some mining explosions is periodic modulation of the spectrum introduced by ripple-firing. A ripple-fired event involves detonation of a number of mining explosions that are often regularly grouped in space and time. Such explosions, known as quarry blasts, have low magnitudes that may be close to those of nuclear explosions that one might monitor under the CTBT. Hence, procedures for detecting ripple firing can serve as bases for discriminating between mining blasts and nuclear explosions or earthquakes. A number of authors have examined various aspects of this problem and have proposed techniques for analyzing these ripple-fired seismic signals (see Baumgardt and Ziegler, 1988, Hedlin et al, 1990, Chapman et al 1992).

The objective of this study is to develop robust statistical procedures for detecting and characterizing the echo structure induced by ripple-firing. This includes developing maximum likelihood in both the time and frequency domains for the purpose of estimating the parameters corresponding to the delay and duration of a ripple-fired signal as well as the spectrum representing source, path and instrument responses that will be convolved with the ripple-fired signal. We use modern model selection techniques based on Akaike's Information Theoretic Criterion (AIC) to determine the number of pulses and their duration.

PRELIMINARY RESEARCH RESULTS

We have developed a time domain approach to the problem of detecting and identifying the echo parameters associated with a ripple-fired event. The underlying model assumes that a sequence of n signal pulses s_t separated by d will produce an observed process of the form

$$y_t = \sum_{j=1}^{n} \alpha_j s_{t-jd} + n_t \tag{1}$$

where n_t is an additive noise. The observed data is, therefore, characterized in terms of a series of scaled replicas of an underlying signal that last for a total *duration* $(n+1)d$ and are observed at delays $d, 2d, \dots, nd$. The signal s_t is an underlying process, assumed to satisfy a low-order autoregressive moving average model of the form

$$s_t - \sum_{k=1}^{p} \phi_k s_{t-k} = w_t - \sum_{k=1}^{q} \theta_k w_{t-k}. \tag{2}$$

The characteristics of the signal process are determined by the p autoregressive parameters $\phi_1, \phi_2, \ldots, \phi_p$, the moving average parameters $\theta_1, \theta_2, \ldots, \theta_q$ and the variance σ^2 of the white noise process w_t. In the usual Box-Jenkins approach (see Shumway, 1988, Chapter 3) the model determined by (2) is identified as $ARMA(p,q)$. It is convenient to rewrite (1) and (2) using the backshift operator B, defined as $Bx_t = x_{t-1}$, and we obtain

$$y_t = \alpha_d(B)s_t + n_t \tag{3}$$

and

$$\phi(B)s_t = \theta(B)w_t, \tag{4}$$

where

$$\alpha_d(B) = 1 + \sum_{j=1}^{n} \alpha_j B^{jd}, \tag{5}$$

$$\phi_(B) = 1 - \sum_{j=1}^{p} \phi_j B^j, \tag{6}$$

and

$$\theta(B) = 1 - \sum_{j=1}^{q} \theta_j B^j. \tag{7}$$

Now, substituting (4) into (3), we obtain

$$y_t = x_t + n_t \tag{8}$$

where

$$\phi(B)x_t = \theta(B)\alpha_d(B)w_t. \tag{9}$$

The observed process y_t is exhibited in (8) as the sum of noise and a multiplicative $ARMA$ process with a seasonal moving average $\alpha_d(B)$ of order n and season d. The process x_t is identified as the multiplicative seasonal $ARMA$, $ARMA(p,q) \times (0,n)_d$. Under this model scenario, the process of detecting a ripple-fired signal is reduced to looking for a multiplicative seasonal moving average component in an ARMA process, where the order corresponds to the number of pulses and the season corresponds to the delay.

The problem of motivating the model defined by (3) and (4) can be approached by appealing to Dargahi-Noubary (1995), who shows that most deterministic source models, including that of Von Seggern and Blandford (1972) can be generated by a stochastic model of the form (4) with $p = 2$ or $p = 3$ corresponding to the ω^2 and ω^3 models respectively if we take the operator as $(1 - \phi B)^p$. We shall not specialize the operator in hopes that the more general form (6) will help to adjust for instrument response and path effects. Furthermore, these effects can be further mitigated by adding general moving average components of order q.

Although the model suggested by (3) and (4) is not exactly $ARMA(p,q) \times (0,n)_d$ because of the additive noise term, it is well known that $ARMA$ signal plus noise models

are equivalent to pure $ARMA$ models with additional moving average components so we are justified in restricting candidates to models of the multiplicative form. What is required is estimating the ARMA parameters $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \sigma^2$ as well as the amplitudes of the reflections $\alpha_1, \ldots, \alpha_n$. This can be done for a fixed p, q, n and d using nonlinear least squares on the residuals, minimizing

$$\hat{\sigma}^2 = T^{-1} \sum_t \hat{w}_t^2, \tag{10}$$

where w_t can be obtained by solving (9) successively, using y_t in place of x_t. There are generally, an unlimited number of values of p, q, n and d possible so we need a method for deciding which values are most reasonable. A common model selection criterion is the corrected form of Akaike's information criterion (see Hurvich and Tsai, 1989),

$$AIC_C = \log \hat{\sigma}^2 + \frac{T + p + q + n}{T - p - q - n - 2}, \tag{11}$$

where we choose the model that minimizes AIC_C. For the large sample used here, this measure gives the same result as the conventional AIC.

In order to illustrate the above methodology, we consider analyzing two contrived events and a single mining explosion from Scandinavia shown in Figures 1, 2 and 3. The contrived events were simulated by generating an $ARMA(2,0)$ process of the form

$$s_t - s_{t-1} + .6 s_{t-2} = w_t \tag{12}$$

for the signal with $\sigma^2 = 1$. This is Equation (2) with $p = 2, \phi_1 = 1, \phi_2 = -.6$. The process was added and delayed using

$$y_t = a_t x_t + n_t, \tag{13}$$

where

$$x_t = s_t + s_{t-8} + s_{t-16} + s_{t-24} + s_{t-32} \tag{14}$$

which is just a delayed signal with $n = 4, d = 8$, i.e. a signal with four pulses separated by 8 points, leading to a duration of 32 points. To make the output resemble an explosion, the $ARMA$ process was modulated by a function of the form $a_t = \theta_1 \exp\{-\theta_2 t\}$ with $\log \theta_1 = -7, \theta_2 = .01$. The second contrived event, shown in Figure 2 adds five pulses at irregular delays and noise n_t with variance $\sigma_n^2 = (.01)^2$. Figure 3 shows a P-phase from a presumed mining explosion recorded at FINESS in Scandinavia and given by Blandford (1993) in his Table 1.

To analyze these three examples, we search for the best $ARMA(p,q) \times (0,n)_d$ model using $AICC$ and the following general guidelines. The operators (6) and (7) that best emulate the convolution of source, path and instrument response, i.e. p and q in the model, are restricted to low orders. In this case, we took $p = 2, 3, q = 0, 1$, corresponding to values suggested by conventional source models and response functions. These low orders all do a respectable job of emulating the observed spectra. In order to limit somewhat the search over the seasonal components n and d, the spectra, log spectra and cepstra were

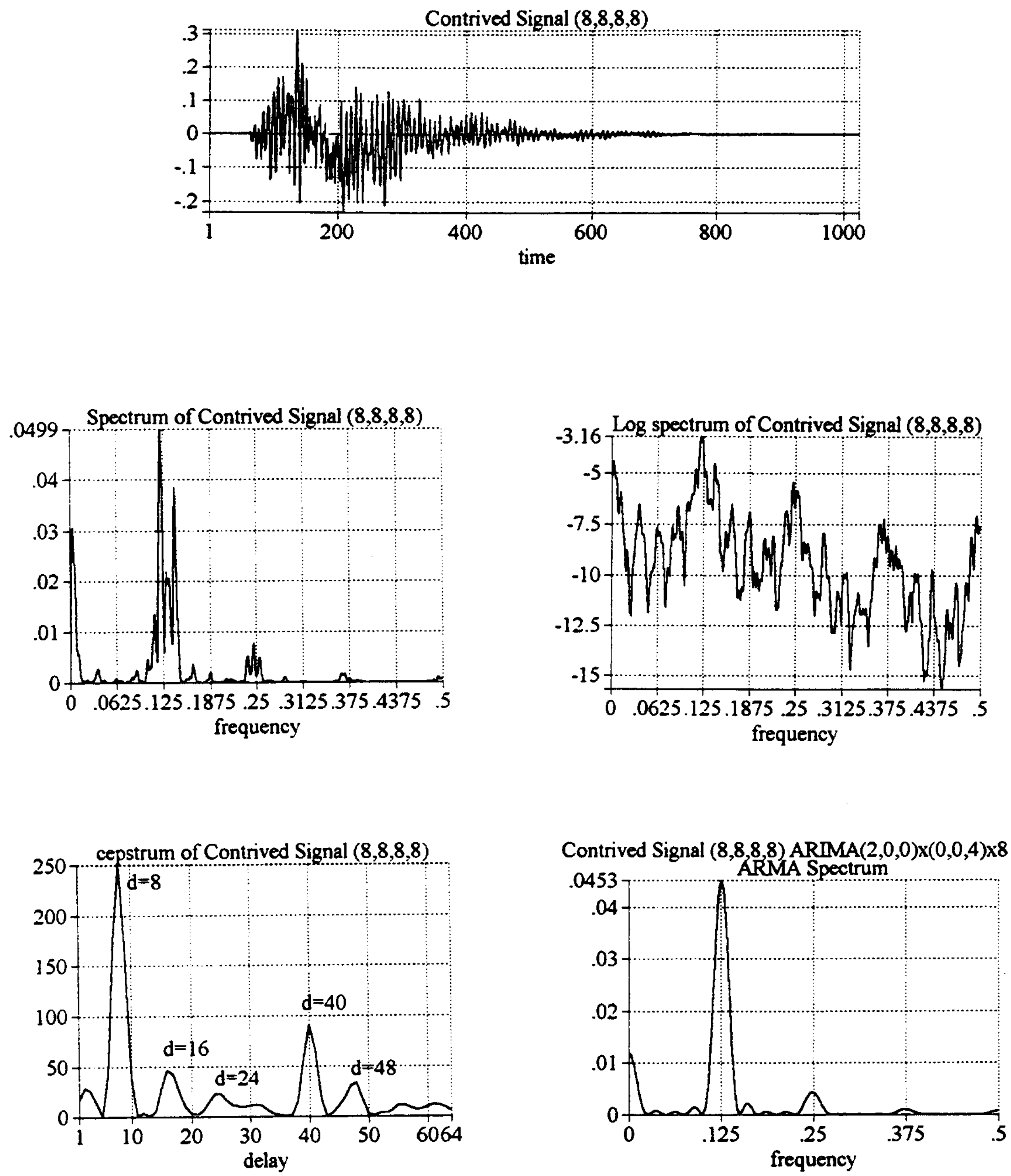

Figure 1. Analysis of a contrived signal composed by modulating an ARMA(2,0) process and delaying using 4 pulses separated by 8 points. Frequency is in cycles/point. Delay is in points.

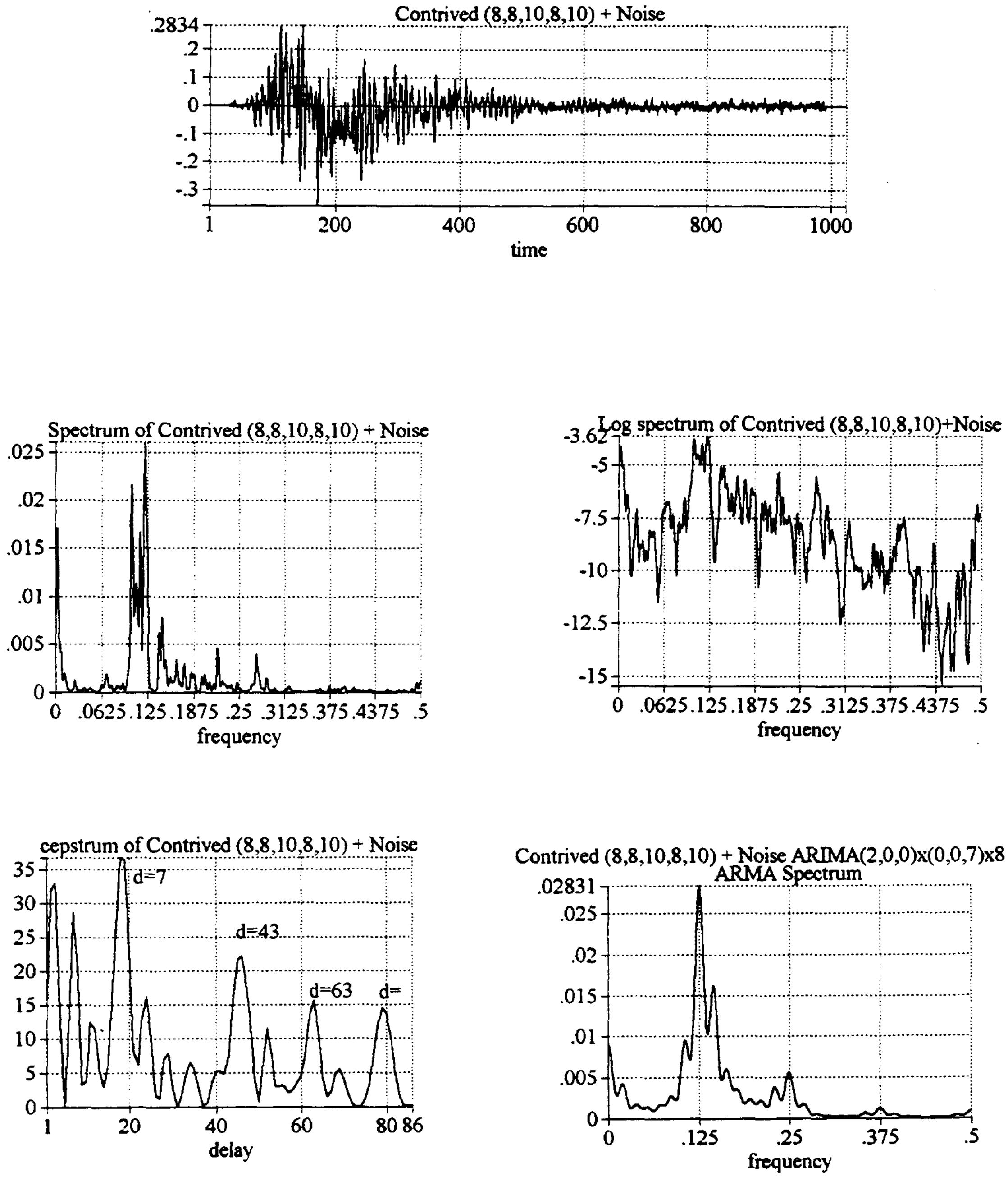

Figure 2. Analysis of a contrived signal composed by modulating an ARMA(2,0) process and delaying using 5 irregularly spaced pulses at intervals 8, 8, 10, 8 and 10 points with additive noise. Frequency is in cycles/point. Delay is in points.

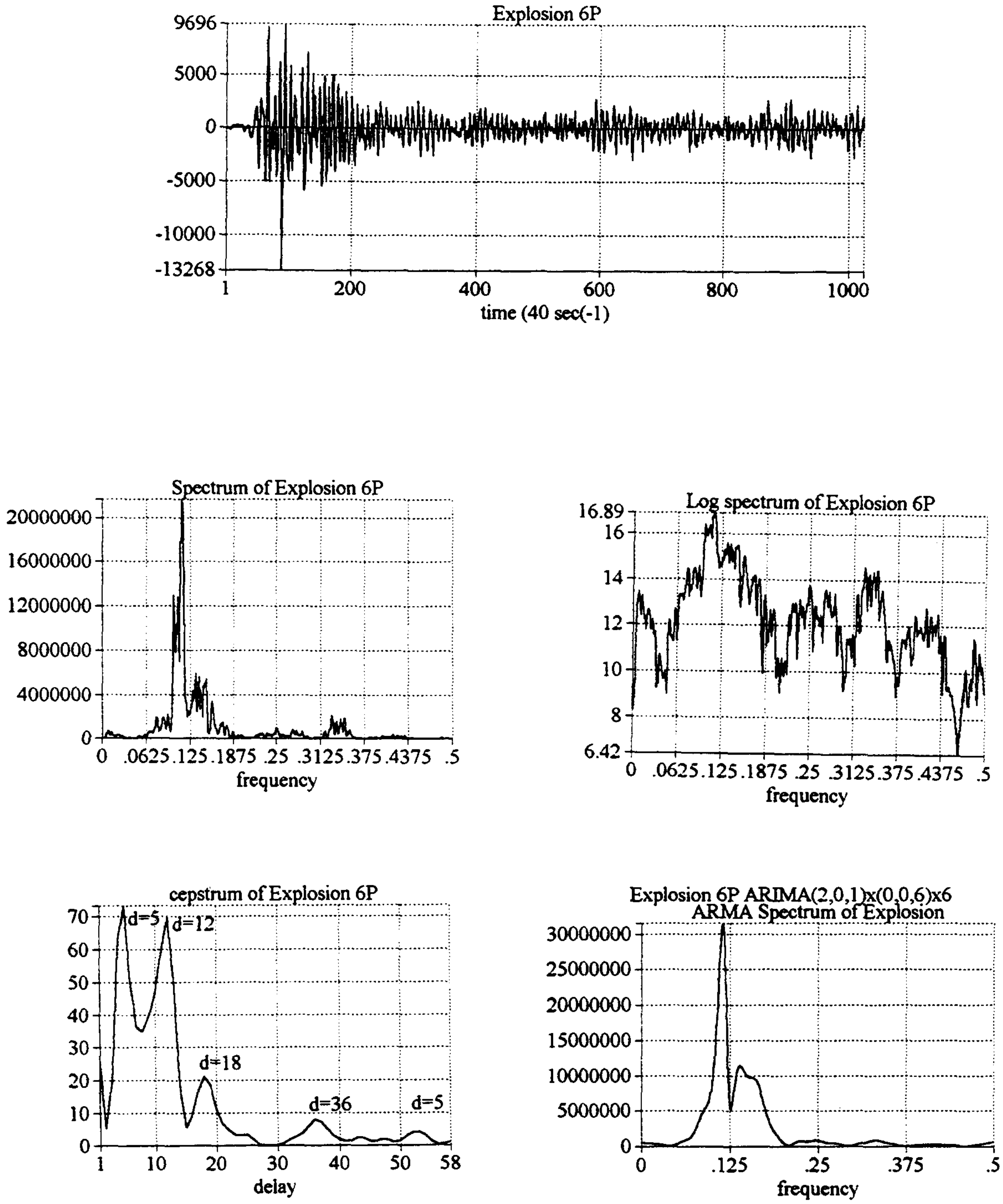

Figure 3. Analysis of a P-phase from a presumed mining explosion recorded at FINESS (F1A1) in Scandinavia (latitude 59.476, longitude 24.1442) dated 12/10/91 with local magnitude 2.59 (see Blandford, 1993, Table 1). Frequency is in cycles per point at 40 sec^{-1} so that the folding frequency corresponds to 20 Hz. Delay is in points or .025 seconds.

computed. The cepstrum, in particular, is helpful in suggesting values for n and d. For example see Bogert et al (1962) who suggested the cepstrum or Baumgardt and Ziegler (1988) who showed, by expansion, the periodicities in the spectrum expected when the spectrum of $x_t = \alpha_d(B)s_t$ is computed. Shumway and McQuarrie (1994) show that when $\alpha_1 = \alpha_2 = \cdots = \alpha_n$, the spectrum of x_t is

$$\begin{aligned} P_x(\omega) &= |\alpha_d(e^{-2\pi i\omega})|^2 P_s(\omega) \\ &= \frac{\sin^2(\pi\omega(n+1)d)}{\sin^2(\pi\omega d)} P_s(\omega), \end{aligned} \tag{15}$$

where $P_s(\omega)$ is the spectrum of the signal. The cepstrum, i.e. $\log P_x(\omega)$, then will be the sum of the logarithm of the signal spectrum and two other components with periods proportional to d, the delay and nd, the duration.

To begin, consider analyzing the first contrived series shown in Figure 1 which we know to be approximately $ARMA(2,0)\times(0,4)_8$ by inspecting the generating equations (12)-(14). This means that there are 4 pulses of length 8 points in the train leading to a duration of $5 \times 8 = 40$ points. Both the spectrum and the log spectrum show periodicities and we can see the two distinct components in the log spectrum. Computing the cepstrum shows main peaks at 8 and 40 corresponding to the delay and duration. To check the seasonal $ARMA$ search procedure, consider searching the models $ARMA(2-3,0-1)\times(0,3-8)_{5-12}$ which indicates looking at $2 \leq p \leq 3, 0 \leq q \leq 1, 3 \leq n \leq 8, 5 \leq d \leq 12$, i.e. low order $ARMA$ models with seasonal components between 5 and 12 and orders between 3 and 8. Table 1 shows the results obtained from the best seven models under the column Figure 1. We see that the minimum corrected AIC is for the correct model.

The best model has the form of (12)-(14) with the estimated parameters substituted, so that

$$s_t - .97s_{t-1} + .60s_{t-2} = w_t$$

becomes the signal model and

$$x_t = s_t + .99s_{t-8} + .87s_{t-16} + .80s_{t-24} + .71s_{t-32}$$

gives the estimated reflection pattern. Note that the coefficients are not all near unity because of the exponential modulation of the x_t in the observed data y_t. Table 1 also gives the next best six models and we note that all but one has the correct delay but that the number of pulses and hence, the duration changes. The spectrum implied by this $ARMA(2,0) \times (0,4)_8$ model is shown in Figure 1 and we note that it agrees quite well with the narrow band spectrum (1 % bandwidth).

Table 1. Model Summaries: $ARMA(p,q) \times (0,n)_d$

Figure	p	q	n	d	AIC_C
1	2	0	4	8	-7.2546
1	2	0	5	8	-7.2543
1	2	0	6	8	-7.2533
1	2	0	8	8	-7.2523
1	2	0	7	8	-7.2513
1	2	0	3	8	-6.8295
1	2	0	6	6	-6.2294
2	2	0	7	8	-5.7271
2	2	0	6	8	-5.7105
2	2	0	5	8	-5.7028
2	2	0	8	6	-5.6993
2	2	0	7	6	-5.6886
2	2	0	6	9	-6.6858
2	2	0	8	9	-5.6848
3	2	1	6	6	14.3649
3	2	1	4	6	14.3706
3	2	1	5	6	14.3714
3	2	1	3	6	14.3773
3	2	1	3	12	14.3830
3	2	1	4	12	14.3848
3	2	1	5	12	14.3865

When some irregularity in the pulse intervals is introduced as in Figure 2, the cepstrum gives quite ambiguous results. From Table 1, note that the seasonal $ARMA$ search indicates a model with 7 pulses separated by 8 when the correct model has 5 pulses; 3 are separated by 8 points and 2 are separated by 10 points. The duration predicted is $(7+1) \times 8 = 64$ points as compared with the known duration of 44 points. The $ARMA$ spectrum shown in Figure 2 is a reasonable version of the narrow band smoothed periodogram. The second best model shows 6 pulses separated by 8 points and an estimated duration of 56 points which seems somewhat closer to the correct model as is the third best model with 5 pulses separated by 8 points. This is a case where the cepstrum gives no insight into the structure but the seasonal $ARMA$ search finds a reasonable model. Again, note that the spacing is easier to estimate than the duration because the number of pulses at a given duration all have about the same value for AIC_C and other durations have generally larger values.

In Figure 3 is shown the P-phase from a single mining explosion along with the spectra, log spectra, cepstrum and $ARMA$ spectrum derived from the best seasonal $ARMA(2,1) \times (0,6)_6$ model. The sampling rate is 40 points per second, leading to a folding frequency of 20 Hz corresponding to .5 cycles per point as shown on the graphs in Figure 3. Table 1 shows that the four best models all have a duration of 6 points or $6/40 = .15$ seconds. A

delay of 150 milliseconds (msec.) may be somewhat long for most mining explosions (see Baumgardt and Ziegler, 1988, Hedlin et al, 1990 or Chapman et al, 1992), but it is clear that this sampling rate may not be high enough to pick up in the range 25-50 msec and we may be looking at aliases. Hopefully, there will be data available at higher sampling rates that can unequivocally sort out reflections in the desired ranges.

RECOMMENDATIONS AND FUTURE PLANS

A high priority for this study is obtaining more multiple-phase (P and S arrivals) explosion and earthquake data sampled at a higher rate, say at 100 points per second. We would also like to obtain high quality signals from an array of elements so as to be able to take advantage of replication, as is done in stacking, practiced by both Baumgardt and Ziegler (1988) and Hedlin et al (1990). For this to work, the increased difficulties inherent in replicated $ARMA$ searching may dictate a more semi-parametric likelihood approach in the frequency domain.

There, we would replace the $ARMA$ model for the spectrum by a more non-parametric version, employing the general form

$$P_y(\omega) = |\alpha_d(e^{-2\pi i\omega})|^2 P_s(\omega) + P_n(\omega), \tag{16}$$

where the signal spectrum $P_s(\omega)$ can be estimated non-parametrically using the Whittle likelihood and assuming that broad-band smoothing will eliminate the duration and spacing ripples. We may also be able to assume that the signal-to-noise ratio is constant over frequency and write

$$P_y(\omega) = P_s(\omega)\left(|\alpha_d(e^{-2\pi i\omega})|^2 + \frac{1}{r^2}\right), \tag{17}$$

where $r^2 = P_s(\omega)/P_n(\omega)$ is the signal to noise ratio. If $P_s(\omega)$ is reasonably constant over a broad band, it can be replaced by the smoothed spectrum over that interval in the Whittle likelihood. Such likelihoods can be treated for arrays as in Der et al (1992).

In the frequency domain, we may also consider the possible of taking into account the amplitude modulation function a_t by considering the Whittle likelihood of the time varying spectrum or sonogram. Dahlhaus (1995) has shown that local averaging over time of the Whittle likelihood can be used to estimate parameters of the modulating function.

A final note relates to the fact that the spacing values d will not be the same and, in fact, may be quite variable. In this case, we must replace jd in (1) by d_j so that

$$y_t = \sum_{j=1}^{n} \alpha_j s_{t-d_j} + n_t, \tag{18}$$

where $d_1, d_2, \ldots, d_n$ form a sequence of unequal delay times. Estimating these as parameters seems intractable because of the high dimensionality of the resulting parameter space. An option that expect to try is letting the delay times be random and possibly uniformly distributed on the integers if one makes no specific assumptions about the time

delays. One may also assume that the random time delays are clustered about the points $d, 2d, 3d, \ldots, nd$ in some specified fashion. The integrated likelihood function can then be computed for any preset combination of parameters using Monte-Carlo methods.

References

Baumgardt, D.R. and K.A. Ziegler (1988). Spectral evidence for source multiplicity in explosions: Application to regional discrimination of earthquakes and explosions. *Bull. Seismolog. Soc. of Amer.*, **78**, 1773-1795.

Blandford, R.R. (1993). Discrimination of earthquakes and explosions at regional distances using complexity. *AFTAC-TR- 93-044*, HQ AFTAC, Patrick AFB, FL.

Bogert, B.P., M.J.R. Healy and J.W. Tukey (1962). The frequency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross cepstrum and saphe cracking. In *Proceedings of a Symposium on Time series Analysis*, ed. M. Rosenblatt. New York: John Wiley.

Chapman, M.C., G.A. Bollinger and M.S. Sibol (1992). Modeling delay-fired explosion spectra and source function deconvolution at regional distances. *Final Report PL-TR-92-2250*, Phillips Laboratory, Directorate of Geophysics, Air Force Materiel Command, Hanscom Air Force Base, MA 01731-3010, *ADA273807.*

Dargahi-Noubary, G.R. (1995). Stochastic modeling and identification of seismic records based on established deterministic formulations. *J. Time Series Analysis* **16** 201-219.

Dahlhaus, R. (1995) Fitting time series models to nonstationary processes. *Beiträge zur Statistik*, Nr. 4, Institut für Angewandte Mathematik, Universität Heidelberg, Heidelberg.

Der, Z.A., A.C. Lees, K.L. McLaughlin and R.H. Shumway (1992). Multichannel deconvolution of short period teleseismic and regional time series. Chapter 9 in *Statistics in the Environmental and Earth Sciences*, A.T. Walden and P. Guttorp (ed., 156-188. Edward Arnold, London.

Hannan, E.J. and P.J. Thomson (1974). Estimating echo times. *Technometrics*, **16**, 77-84.

Hedlin, M.A.H., J.B. Minster and J.A. Orcutt (1990). An automatic means to discriminate between earthquakes and quarry blasts. *Bull. Seismolog. Soc. Amer*, **80**, 2143-2160.

Hurvich, C.M. and C.L. Tsai (1989). Regression and time series model selection in small samples. *Biometrika* **76** 297-307.

Shumway, R.H. (1988). *Applied Statistical Time Series Analysis*, Chapter 5. Englewood Cliffs: Prentice-Hall.

Shumway, R.H. and McQuarrie, A.D.R. (1994). Statistical discrimination studies for nuclear test verification. Final Report, PL-TR-94-2283, Phillips Laboratory, Directorate of Geophysics, Air Force Materiel Command, Hanscom AFB, MA 01731-3010. ADA293572

Von Seggern, D. and R. Blandford (1972). Source time functions and spectra of underground nuclear explosions, *Geophysical J. R. Astro. Soc.*, **31**, 83-97.